NETROLOGIC, Inc.
5080 Shoreham Place
Suite #201
San Diego, CA 92122

ONRA
Interim Report #2

AUG 1990

Item 0001AB

Contract #N00014-90-C-0124

SCOPE OF REPORT

This report covers algorithm development and theoretical work performed during the months of July and August.

ALGORITHM DEVELOPMENT

WORK TO PRESENT. During July and August we continued our investigation of correlation methods for coding images and began to investigate wavelet transforms. The work included six areas of interest:

1) Lossless compression;
2) Alternative tile classification schemes;
3) Elimination of compression artifacts;
4) Effects of image size on fidelity and compression;
5) Image preprocessing; AND
6) Methods to improve tiling fidelity.

While the compression ratios attainable with lossless compression schemes are far lower than those obtained with lossy schemes, the utility of lossless schemes is higher for many applications. We investigated briefly the idea of compressing the difference between an original image and a highly compressed version of the image. The results were encouraging, though not spectacular. Compression of our standard test images ranged from 1.3 to 1.9. We plan to investigate other lossy techniques, specifically Fourier transforms, to obtain more comparisons.

The second area of interest, using alternative classification schemes for tiles, is somewhat technical. The enclosed preprint has further details. The domain tiles used to encode an image are classified in order to speed the search needed to find a 'good' domain tile. Using a classification scheme generally results in poorer fidelity, because it is not guaranteed that the optimal domain will be found in the class searched. We investigated several methods based on correlation methods and moments. The results are not definitive, being better in some features and worse in others as compared with the present scheme.

One weakness of the current compression scheme is the appearance of artifacts. In order to eliminate these we attempted to postprocess the image by smoothing along the boundaries

of the range tiles. This was successful, often resulting in an even lower rms error. Of the attempts to remove compression artifacts, this was the most successful. Figure 1 shows a 512x512 pixel 8bpp image of Lena; figure 2 shows a smoothed version the decoded image of Lena at a compression of 38.7:1; figure 3 shows this same image before smoothing.

Another area of interest is the effect of image size on fidelity and compression. Somewhat unexpectedly, encoding larger images results in significantly better compression and fidelity. This observation was uniformly true for all the test images we experimented with. Figures 2 and 3 can be compared with the images in the preprint to show the large improvement in compression for a comparable fidelity.

In an effort to improve the ability of the encoding scheme in encoding images, we investigated two image preprocessing methods. The first, based on Wigner transforms seems to be initially disappointing. The second method using a wavelet transform, is potentially useful. We discuss more of the technical details in succeeding sections. Figure 4 shows the result of wavelet transforms applied to an image.

Finally, we are attempting to improve the fidelity of a given tiling by using linear combinations of domain tiles. Initial results are very encouraging, having greatly reduced artifacts. Using more than one domain tile significantly reduces the error, but decreases the overall compression. Without careful classification the search time needed to find even a moderate approximation to an optimal fit is prohibitive. Our current approach to using several domain tiles in an encoding is to encode a range optimally and then encode the resulting error. We are investigating alternative methods which may yield a better approximation to the optimal linear combination without increasing the computation time excessively. Figure 5 shows the reconstruction of an image for which the encoding is two domain tiles to each range tile.

PUBLICATIONS. Dr. Yuval Fisher has submitted a paper for publication in the IEEE Transactions on Acoustics, Speech and Signal Processing.

ANTICIPATED WORK.

MODIFICATION OF TILING. Work so far has concentrated on tilings which are regular, rectangular and non-overlapping. This limits the choice of affine transformations and does

not give sufficient flexibility with respect to the local spatial frequency content of the image. We will investigate ways to choose relatively larger domains where the image is not rapidly changing and smaller domains where the image is varying rapidly. We will investigate these methods in conjunction with the Fourier transform.

TRANSFORM METHODS. One efficient way of determining the local frequency content of the image is through transforms related to the Fourier transform. These, in order of increasing generality, are the Fourier transform, the Wigner transform, and wave packet transforms. The Wigner transform is given by the integral

$$W(f) = \int e^{-2\pi i \xi p} \, f(x + \tfrac{1}{2} p) \, f(x - \tfrac{1}{2} p) dp \, ,$$

and a generalized wave packet transform is given by the integral

$$P_{\phi}^{f}(p, q, t) = \int e^{-\pi i p q + 2\pi i p x} \, t^{1/2} \, \phi[t^{1/2}(x - q)] \, f(x) \, dp \, ,$$

where $\phi(x)$ is a generalized Gaussian (see below). A little consideration of the definition of $W(f)(x, \xi)$ will show that this transform gives the local frequency content of the function f in a neighborhood of x. The expression for $P_{\phi}^{f}(p, q, t)$ is the inverse Fourier transform for W(f) if we set $t = 1$ and $\phi = f$. Details may be found in Folland's 1989 monograph, pp 56-63 and 142-169.

WAVELET TRANSFORM METHODS. During the next two months we will investigate relationships between fractal methods and wavelet transforms. The objective of this phase of the study is to improve the choice of affine iterated function systems. Our rationale for doing this is that our current image compression scheme does not preserve edges well. Because edges contribute heavily to human perception of image quality, it is important to amend this weakness. An appropriate wavelet transform yields a powerful tool to extract edge information from an image on a variety of scale lengths. Figure 4 shows edges extracted from a wavelet transform of "Lena" at different scale lengths. Edges correspond

to zero crossings of a wavelet transform of an image and the image may be reconstructed, by standard methods, from such zero crossings.

Edge information may be used in conjunction with fractal methods in various ways. One approach is to force the coding method to preserve edges by using a error measure which weights the edge error more heavily. (Recall that our encoding scheme chooses transforms on the basis of an integrated error--see the enclosed preprint, page 4.) An alternative approach is to regard the edges as one-dimensional fractal curves. The edges themselves may be coded as iterated transforms, using a 1-dimensional analog of the 2-dimensional method used for images. We will investigate both of these approaches during September and October.

THEORETICAL INVESTIGATIONS

WORK TO PRESENT. Up to the present we investigated the behavior of generalized Gaussian functions under affine transformations. A generalized Gaussian is given by a function:

$$\phi_{A,b,d}(x) = \exp(x^t Ax + b^t x + d),$$

where x is a vector in the plane, A is a symmetric, negative definite matrix, and b is a fixed vector. There is also a complex-valued form, in which the entries of A and b are complex numbers and the real part of A is required to be negative definite. Figures 6 through 9 show sine and cosine components of generalized Gaussians. The set of generalized Gaussians is preserved under non-singular affine transformations and the orbits (up to a complex factor) are specified by the cogredience class of A. According to a theorem of Sylvester, any two real symmetric matrices are cogredient if they have the same rank and the same signature (Jacobson, 1953). The closure property may be seen by a simple calculation. If $x \rightarrow Bx + c$ is an affine transform $\mathbf{A}$, then $\phi(x)$ is transformed into

$$\mathbf{A}(\phi_{A,b,d})(Bx + c)$$
$$= \exp(x^t B^t ABx + c^t Ax + x^t Ac + b^t Bx + c^t Ac + b^t c + d).$$

The latter function is of the form $\phi_{A',b',d'}$, where $A' = B^t AB$, $b = 2c^t A + b^t B$ and $d' = c^t Ac + b^t c + d$. Because A is symmetric and negative definite the equation

$$2c^t A + b^t B = e^t$$

has a unique solution c. Therefore by proper choice of affine transforms, the quadratic part, $x^t Ax$, is specified up to cogredience class, the linear part, $b^t x$, is arbitrary and we are left with a constant factor which cannot be specified.

The objective of our planned work is to clarify the relationship between image coding by wavelets and image coding by iterated function systems. The invariance of the set of generalized Gaussian functions under affine transformations indicates that the analog of a

fractal should be a sum of Gaussians which is preserved under an iterated function system. In particular, if

$$f = \sum_i \phi_{A_i, b_i, d_i},$$

then we look for a system of affine transformations $\{\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_N\}$ with weights $\{w_1, w_2, \cdots, w_N\}$ such that

$$f = \sum_{i=1}^{N} w_i (\mathbf{A}_i)^{-1}(f) = \sum_i \phi_{A'_i, b'_i, d'_i}.$$

Because we require

$$\sum_i \phi_{A'_i, b'_i, d'_i} = \sum_i \phi_{A_i, b_i, d_i},$$

the summations need not be unique. We plan to investigate the problem of finding decompositions which are linked by systems of affine transformation.

This work has shown some connections between generalized harmonic analysis and the theory of fractals, in particular through representations of the extended metaplectic representation and the inhomogeneous symplectic group. The metaplectic representation is a representation of the symplectic group (the group of $2n \times 2n$ matrices which preserve the symplectic form $[(\mathbf{p}, \mathbf{q}), (\mathbf{p}', \mathbf{q}')] = \mathbf{pq}' - \mathbf{p}'\mathbf{q}$, on vectors with $2n$ components. The symplectic group is generated by matrices

$$\begin{pmatrix} A & 0 \\ 0 & A^{*-1} \end{pmatrix}, \begin{pmatrix} I & 0 \\ C & I \end{pmatrix} \text{ and } \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$$

The action is given, up to sign, by

$$\mu\left[\begin{pmatrix} A & 0 \\ 0 & A^{*-1} \end{pmatrix}\right] f(x) = (\det{}^{-1/2} A) f(A^{-1}x)$$

$$\mu\left[\begin{pmatrix} I & 0 \\ C & I \end{pmatrix}\right] f(x) = \pm e^{-\pi i C x} f(x)$$

$$\mu\left[\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}\right] f(x) = i^{-n/2} \mathcal{F}^{-1} f(x),$$

where $\mathcal{F}$ is the Fourier transform. The extended metaplectic representation is a representation of the semidirect product of the Heisenberg and symplectic groups. The semidirect product is given as pairs of operators $(X, \mathcal{A})$, where X is in the Heisenberg group and $\mathcal{A}$ is in the symplectic group. The group product is given by $(X, \mathcal{A})(X', \mathcal{A}') = (X(\mathcal{A}X'), \mathcal{A}\mathcal{A}')$, and the representation is given by $\omega(X, \mathcal{A}) = \rho(X)\mu(\mathcal{A})$. The Heisenberg group acts as follows for $X = (p, q, t)$:

$$Xf(x) = e^{2\pi i t} e^{2\pi i q x + \pi i p q} f(x + p).$$

Because the product in the Heisenberg group is given by

$$(p, q, t)(p', q', t') = (p + p', q + q', t + t' + \tfrac{1}{2}(pq' - qp')),$$

it can be seen that the extended metaplectic representation contains the usual action of the affine group.

The inhomogeneous symplectic group is another extension of the symplectic group, which has a more obvious relationship to the affine group. In this case the semidirect product is the product of R^{2n} with the symplectic group, and the group multiplication is

$$(w, \mathcal{A})(w', \mathcal{A}') = (w + \mathcal{A}w', \mathcal{A}\mathcal{A}').$$

The representation, as before, is given by $\omega(X, \mathcal{A}) = \rho(X)\mu(\mathcal{A})$. This is a projective representation, and the extension of this representation by a representation of the circle group is the extended metaplectic representation.

OTHER CONNECTIONS BETWEEN GENERALIZED GAUSSIANS AND AFFINE TRANSFORMATIONS. It is easy to show that the set of generalized Gaussians is preserved

under the convolution product. This may be derived from results on the oscillator semigroup (Folland, page 231) or calculated directly. Furthermore, affine transformations preserve the convolution product, up to a constant factor. If we write

$$\gamma_{[A,x_0,c]} = \exp[(x - x_0)^t A (x - x_0) + c]$$

where, as before the real part of A is negative definite, then

$$\mathbf{M}\,\gamma_{[A,x_0,c]} = \gamma_{[M^t A M,\ M^{-1}(x_0 - b),\ c]},$$

where $\mathbf{M}(x) = M(x) + b$, and

$$\gamma_{[A,x_0,c]} * \gamma_{[B,y_0,c]} = \gamma_{[A^t(A+B)^{-1}A + A,\ x_0 + y_0,\ c']}$$

where $c' = c - 1/2 \log(\det(A + B))$ and $*$ represents the convolution produce.

This can be used to derive relationships between wavelet transforms and fractal codings.

AFFINE TRANSFORMATIONS AND THETA FUNCTIONS. One definition for the theta function is given via a generalization of the Fourier transform. We let ${}_1x$ be a vector in $\mathbf{R}^k$ and ${}_j x$ stand for a vector in the space ${}_j \mathbf{R}^k$ of homogeneous polynomials of degree j in $\mathbf{R}^k$. There is an obvious inner product in this space, so we can define the nonlinear Fourier transform

$$F(\tau^1, \ldots, \tau^J) = \int f({}_1x)\exp[i\tau^1 \cdot {}_1x, \ \ldots, \ i\tau^J \cdot {}_J x]\, dx .$$

If we take f to be $\sum \delta_1 n$, where $\{{}_1 n\}$ is the lattice of vectors with integer coordinates in $\mathbf{R}^k$ and J = 2, then we get the function $\theta_3(x, z)$. Furthermore,

$$\int f(\mathbf{M}_1 x)\exp[\, i\,\tau^1{}_1 x,\ \ldots,\ i\,\tau^J{}_J x]\, dx =$$

$$\mathrm{Det}(M)^{-1} \int f({}_1 x)\, \exp[\, i\,\tau^1 \cdot \mathbf{M}^{-1}{}_1 x,\ \ldots,\ i_J(\mathbf{M}^{-1})_J x]\, dx =$$

$$\mathrm{Det}(M)^{-1} \int f({}_1 x)\, \exp\{i\,[\mathbf{M}^{-1}]^*\tau^1{}_1 x,\ \ldots,\ i\,[_J(\mathbf{M}^{-1})]^*\tau^J{}_J x\}\, dx =$$

$$\mathrm{Det}(M)^{-1}\, F(\,[\mathbf{M}^{-1}]^*\tau^1,\ \ldots,\ [_J(\mathbf{M}^{-1})]^*\tau^J\,)$$

where * denotes the adjoint and the pre-subscript denotes the appropriate symmetric product. This shows that approximate symmetries under the affine group of the function f can be represented as approximate symmetries of the affine group acting on the nonlinear Fourier transform. This indicates that searching zero-crossings should be a good way to find appropriate transforms to code an image.

REFERENCES:

N. Jacobson *Abstract Algebra, Volume 1*, Van Nostrand, New York, 1965.

R. Folland *Harmonic Analysis in Phase Space*, Princeton University Press, Princeton, NJ, 1988.

Figure 1.

The original Lena at size 512 x 512 and 8bpp

Figure 2.

A smoothed 512 x 512 encoded Lena.
Compression: 38.7
RMS error 8.23 (29.82db signal to noise ratio)
1666 affine transformations.

Figure 3.

An unsmoothed 512 x 512 encoded Lena.
Compression: 38.7
RMS error 8.81 (29.23db signal to noise ratio)
1666 affine transformations.

Figure 4.

A wavelet transform of a 256 x 256 8pbb version of Lena resulting in the extraction of edge information.

Figure 5.

An unsmoothed 256 x 256 Lena encoded using linear combinations of domain tiles. This images shows almost no boxy artifacts.

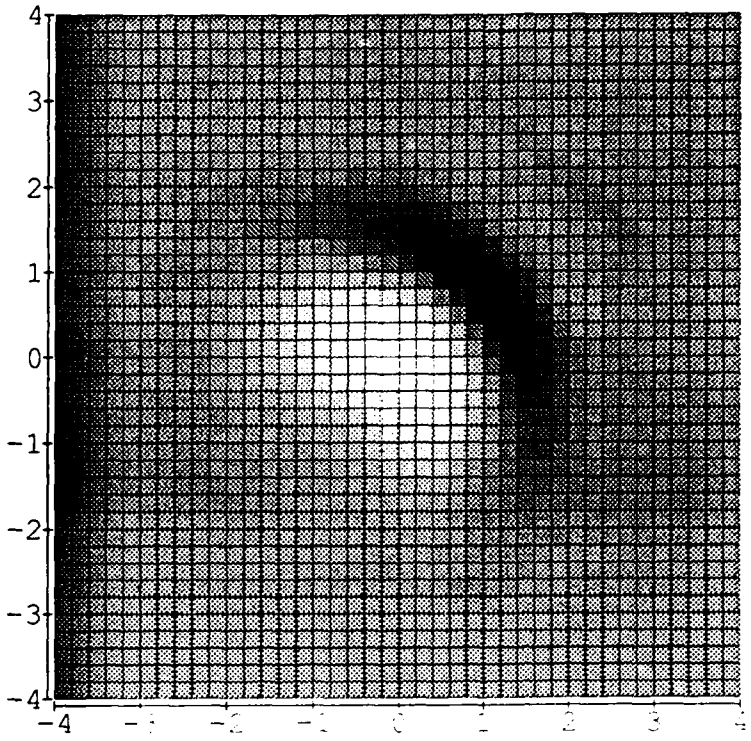

Figure 6a. Elliptic Gabor function, cosine component

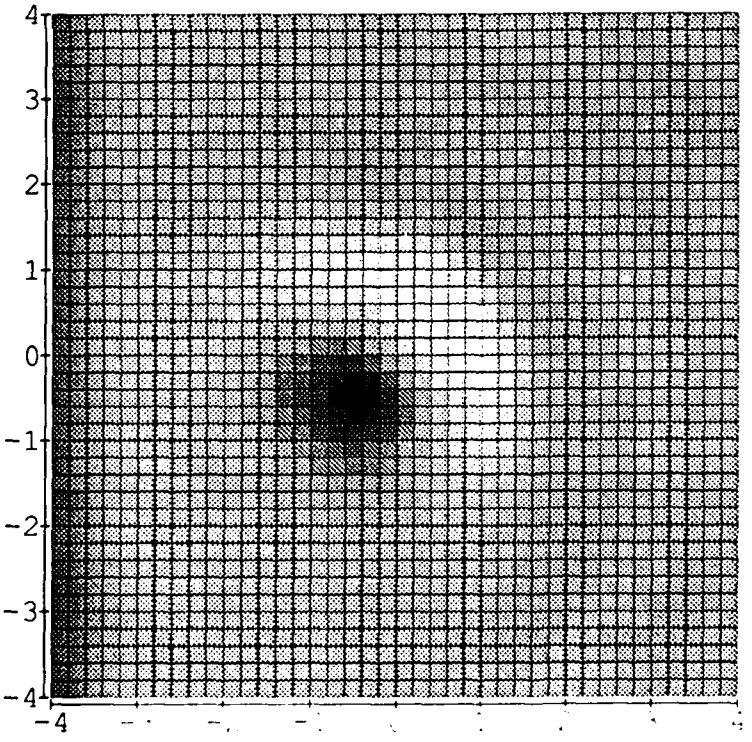

Figure 6b. Elliptic Gabor function, sine component

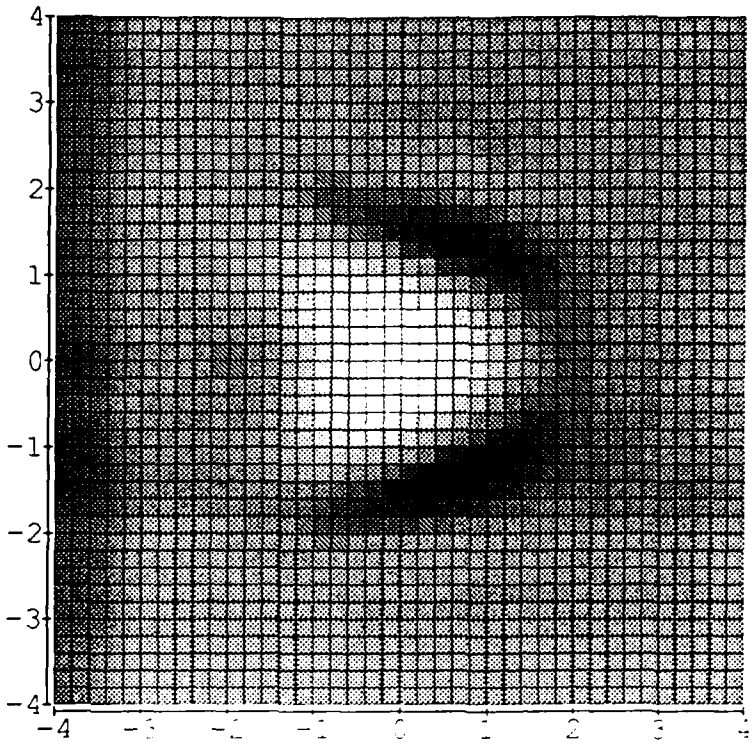

Figure 7a. Parabolic Gabor function, cosine component

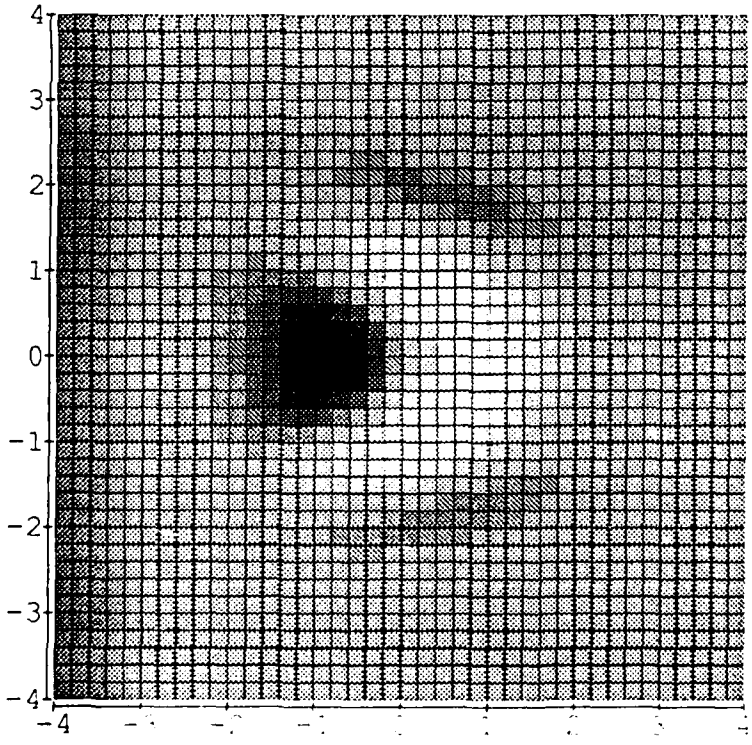

Figure 7b. Parabolic Gabor function, sine component

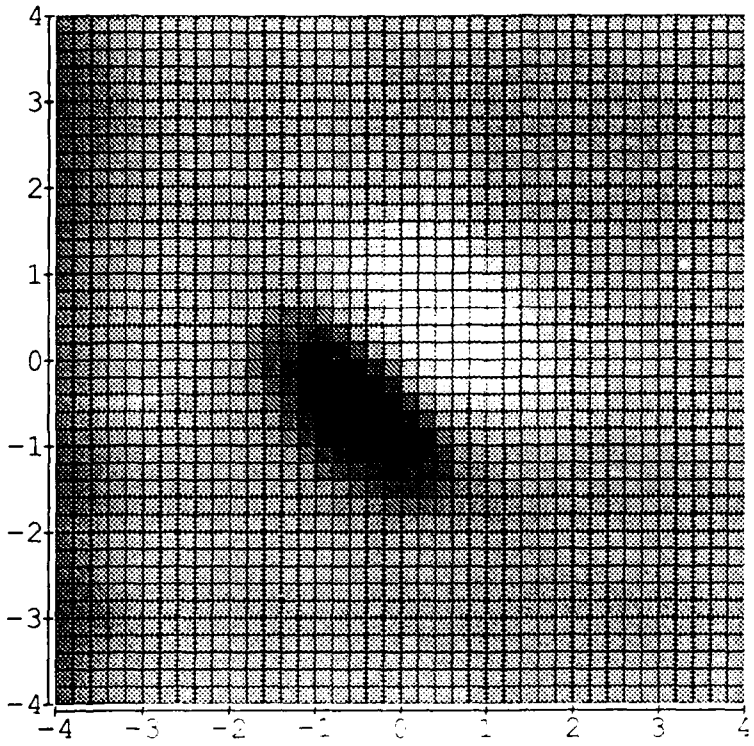

Figure 8a. Linear Gabor function, sine component

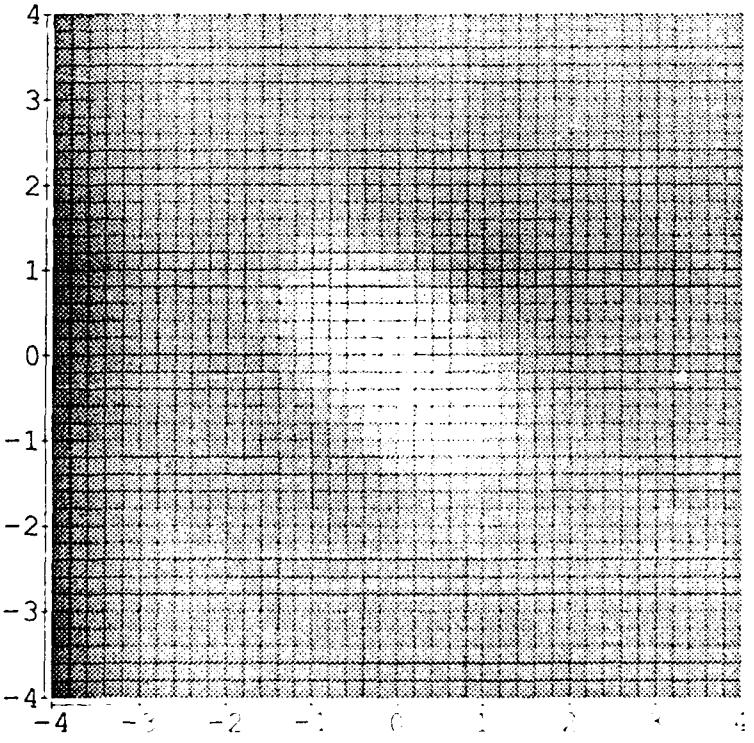

Figure 8b. Linear Gabor function, cosine component

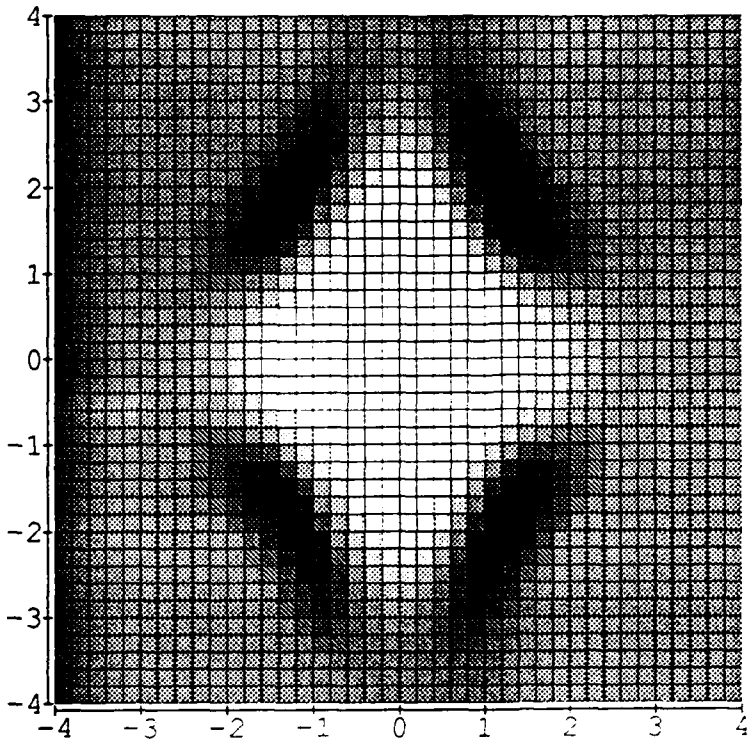

Figure 9b. Hyperbolic Gabor function, cosine component

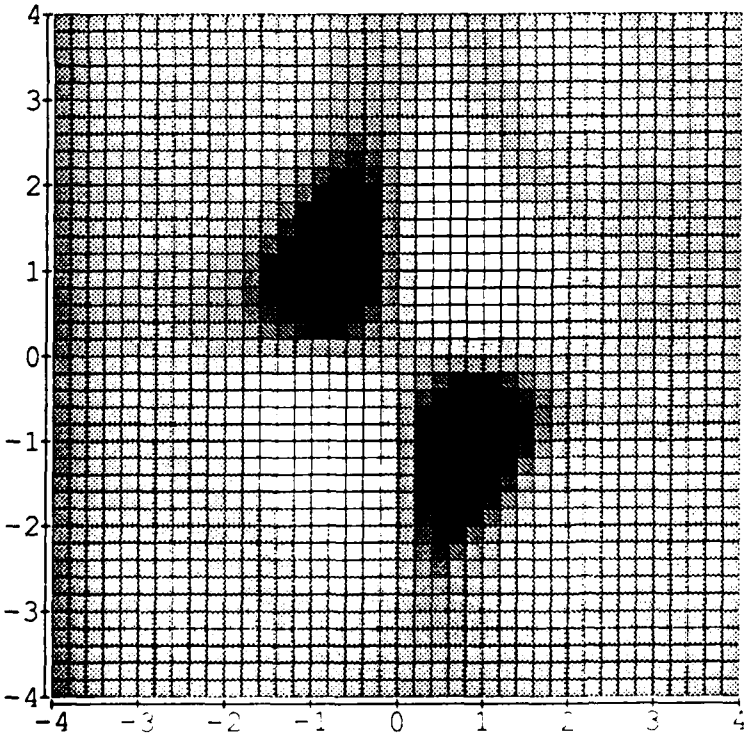

Figure 8b. Hyperbolic Gabor function, sine component